Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 10869–10873, November 1995
Biochemistry

Optimization of rates of protein folding: The nucleation–condensation mechanism and its implications

(denatured state/fragments/intermediates/evolution/computer simulation)

ALAN R. FERSHT

Medical Research Council Unit for Protein Function and Design and Cambridge Center for Protein Engineering, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, United Kingdom

Contributed by Alan R. Fersht, August 25, 1995

ABSTRACT **Small, single-module proteins that fold in a single cooperative step may be paradigms for understanding early events in protein-folding pathways generally. Recent experimental studies of the 64-residue chymotrypsin inhibitor 2 (CI2) support a nucleation mechanism for folding, as do some computer simulations. CI2 has a nucleation site that develops only in the transition state for folding. The nucleus is composed of a set of adjacent residues (an α-helix), stabilized by long-range interactions that are formed as the rest of the protein collapses around it. A simple analysis of the optimization of the rate of protein folding predicts that rates are highest when the denatured state has little residual structure under physiological conditions and no intermediates accumulate. This implies that any potential nucleation site that is composed mainly of adjacent residues should be just weakly populated in the denatured state and become structured only in a high-energy intermediate or transition state when it is stabilized by interactions elsewhere in the protein. Hierarchical mechanisms of folding in which stable elements of structure accrete are unfavorable. The nucleation–condensation mechanism of CI2 fulfills the criteria for fast folding. On the other hand, stable intermediates do form in the folding of more complex proteins, and this may be an unavoidable consequence of increasing size and nucleation at more than one site.**

A purely random search of an unfolded polypeptide for its native conformation would have to explore an astronomical number of conformations (1). Many mechanisms have been proposed to simplify the process. A major class of mechanism invokes the formation of a protein in stages, with the hierarchical accretion of stable, smaller elements of structure (2–7). Alternatively, it has been proposed that a protein can rearrange from a collapsed structure (8, 9). The field of protein folding has accordingly been dominated for many years by the notion that the formation of stable folding intermediates is an essential element of protein-folding pathways. Metastable products, sometimes called molten globules, can frequently be isolated and have been assumed to be general reaction intermediates (10–12). Intermediates undoubtedly occur on folding pathways, and the presence of intermediates is useful since the characterization of successive intermediates along a pathway is the best way of establishing a pathway (13). There is a fundamental problem about stable intermediates, however, in that the accumulation of an intermediate on a pathway is not necessarily desirable for the optimization of the rate of the pathway (14–17).

The earlier evidence about intermediates in folding tended to come from experiments on large proteins (>150 residues). In the last 5 years, however, it has become apparent that an increasing number of small proteins fold rapidly by simple two-state kinetics, without the accumulation of an intermediate and with only one kinetically important transition state (17–25). Understanding the mechanism of folding of small proteins containing a single module of structure may be the key to understanding the initial events in protein folding in general, and it has been suggested that their folding affords models for the folding of individual modules of structure in larger proteins (26).

Analysis of the microscopic mechanisms of folding needs evidence at the level of individual residues. The key structures to be analyzed on the folding pathway of those small proteins are the denatured state and the transition state. Although the structures of such transition states cannot be directly studied, methods have been devised to analyze them at microscopic levels by kinetic and equilibrium measurements on suitable mutants [the "protein engineering method" (27–29)]. At the same time, computer simulation methods have been developed to analyze the folding of simplified models of small proteins. Lattice models, in which simplified amino acid residues that are represented as structureless beads are "folded" on a cubic lattice, are especially applicable to simulating the early events in folding (9, 30–38). The lattice models make predictions about the topology of the polypeptide chain during folding but not about the packing of individual side chains. However, the results depend crucially on the potentials used, and the conclusions must be validated by experiment. The protein-engineering procedure gives the interaction energies of side chains in the transition state and shows whether they are in a denatured, fully folded, or partly formed conformation.

Experiment and theory have converged on the importance of nucleation mechanisms (26, 34, 39). I wish to assess some implications of these results with a simple thermodynamic model.

Folding Mechanism of Chymotrypsin Inhibitor 2 (CI2): Nucleation–Condensation

The truncated form of CI2 from barley consists of 64 amino acid residues in a single, unbranched polypeptide chain. Its secondary structure comprises a single α-helix, spanning residues 12–24 and a mixed parallel and anti-parallel β-sheet. The structures of a series of peptide fragments of CI2 of increasing length from the N terminus have been examined in detail (40, 64). Only 3% or less of the sequence containing the helix in the native structure takes up a helical structure in shorter peptides, and the helix is not stabilized until more than 50 residues are present and long-range interactions are made. The remaining sequences in fragments of CI2 are largely unstructured. Under physiological conditions, the denatured protein is highly solvated and devoid of fixed structure, as judged by relative changes in specific heat for the kinetics and equilibria of folding (41) and its titration properties (65).

Unlike all previously studied proteins, CI2 was found in 1991 to fold and unfold via simple two-state kinetics without the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: CI2, chymotrypsin inhibitor 2.

accumulation of any stable intermediate (18, 19), and the folding was fast ($t_{1/2}$ = 14 ms at 25°C). The structure of the single rate-determining transition state for the folding in water is built around a nucleation site that consists of the N-terminal part of the α-helix (residues 12–21) and some extended interactions (ref. 26; L. S. Itzhaki, D. E. Otzen, and A.R.F., unpublished results). The helix is in the process of being formed in the transition state and is stabilized by interactions elsewhere in the protein, especially those of the side chain of Ala-16 with Leu-49 and Ile-57 (L. S. Itzhaki, D. E. Otzen, and A.R.F., unpublished results). The rest of the protein is only weakly structured and is in the process of collapsing around the nucleation site as the nucleation site itself is being formed. I prefer the description "nucleation–condensation" for this mechanism rather than "nucleation–growth" that is usually used for homogeneous phase transitions.

Abkevich *et al.* (34) have shown independently from lattice model simulations that nucleation–growth (or –condensation), popularized by Wetlaufer (42), should be an efficient mechanism of folding and have discussed some of its implications and history (see also ref. 39). The complementary analysis below shows why a nucleation site composed of distant, as well as a cluster of adjacent residues, is advantageous; why residual structure in fragments is unfavorable; and why the nucleation–condensation mechanism, which follows simple two-state kinetics without the accumulation of an intermediate, is a highly evolved and efficient mechanism for folding. The basic principles may be exemplified by the kinetics of enzymes since their reactions are well studied and many enzymes have been subjected to evolutionary pressures to maximize their rates of reactions.

Accumulation of Intermediates Slows Down Enzymatic Reactions

In the early 1970s, it was widely thought that enzymes should bind their physiological substrates tightly. For example, I was taught in 1969 that there was a paradox about carbonic anhydrase because the physiological concentrations of its substrates, carbon dioxide and bicarbonate, are some ten times lower than the values of their K_ms in the enzymatic reactions, so that substrate is bound only to a small fraction of the enzyme. It was expected that an enzyme should be saturated with its specific substrates under physiological conditions, since the fraction of the enzyme that was bereft of substrate would be unproductive. This is a misconception because optimal rates are obtained when all intermediates on the reaction pathway are at a higher energy level than the starting materials and so do not accumulate but exist at low concentrations (14). The reasoning behind this has been discussed in detail (43). It has been accepted since the proposals by Haldane in 1930 (44) and Pauling in 1946 (45) that the active site of an enzyme has evolved to be complementary in structure to the transition state ($S^‡$) of its substrate (S). This fixes the energy level of the enzyme–transition state complex ($E.S^‡$). On the left of Fig. 1 is the free energy profile for an enzyme-catalyzed reaction for which the substrate is bound sufficiently tightly to keep the enzyme–substrate complex (E.S) at a lower energy level than the free enzyme and substrate at its physiological concentration—i.e., $[S]_{phys} > K_m$—thus, making E.S accumulate. On the right of Fig. 1 is the energy profile for the same reaction, but the enzyme has evolved to bind the same substrate weakly so that $[S]_{phys} < K_m$ and, therefore, E.S is present only at low concentration. Since the ground state for the reaction is the E.S complex, the activation energy of the reaction is higher for the tightly bound substrate than for the weakly bound substrate, where the ground state of the reaction is the free reagents. A low-energy enzyme–substrate complex is a "thermodynamic pit," out of which the reaction has to climb. A corollary of this argument is that, in efficient enzymes, intermediates formed after the E.S complex should not accumulate under physiological conditions—e.g., the acyl-enzyme in serine protease-catalyzed reactions). There is, thus, evolutionary pressure on enzymes that have to function at maximal rate to bind their substrates weakly and not accumulate intermediates, besides binding the transition states tightly.

Whether any one real enzyme has evolved to bind its substrates sufficiently weakly depends on whether evolutionary pressure had to maximize its catalytic rate (since some enzymes have their rate subservient to other functions) and, if so, whether it could discriminate sufficiently well between the difference in structures of its particular substrates and its transition state. The constraints will differ from reaction to reaction. A survey of the physiological concentrations of substrates and the associated values of K_m in glycolysis revealed that K_ms are indeed generally greater than $[S]_{phys}$ (43) and, therefore, the behavior of carbonic anhydrase is not exceptional. Those enzymes' structures have thus responded to evolutionary pressure to maximize rate.

This is not to deny the occurrence of intermediates in enzymatic reactions, but they tend not to accumulate under physiological conditions and sometimes do not accumulate at all with natural substrates under any conditions. Experimentalists frequently detect intermediates at unnaturally high concentrations of substrates, under extreme conditions, or by using unnatural substrates.

Similar energy profiles can be drawn for the folding of proteins, but there are several points to be considered before extending the analysis of enzyme catalysis directly to protein-folding pathways. First, there is the question of evolutionary pressure. It is advantageous to increase the efficiency of enzyme catalysis to the maximum since the cell has to produce less enzyme. But, proteins have to fold merely at an adequately fast rate. Similarly, it has been possible to increase the stability of barnase by a series of point mutations without decreasing its activity (46), suggesting that stability evolves only as far as necessary. The question is whether there is evolutionary pressure to maximize the rate of protein folding and, if there is, whether

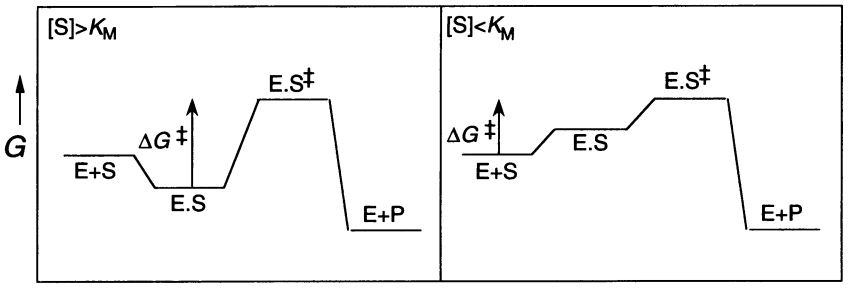

FIG. 1. Energy levels for enzyme catalysis (adapted from refs. 14 and 43). (*Left*) E.S accumulates since $[S]_{phys} > K_m$ ($[S]_{phys}$ = physiological concentration of substrate, standard state for S is $[S]_{phys}$). (*Right*) E.S does not accumulate since $K_m > [S]_{phys}$. The free energy of activation is lower on the right when the intermediate state is at higher energy.

formation of their structures follows simple rules similar to those governing optimization of catalysis.

Energy Diagrams for Protein Folding

The thermodynamic/kinetic analysis used for enzyme catalysis has to be modified because we are dealing with the low energy of individual noncovalent bonds. The energy profiles have to be modified as in Fig. 2. The nominally unfolded state, U, of a protein under denaturing conditions is not a discrete species but is an ensemble of conformations. The same applies to the denatured state under physiological conditions, D_{phys}, which is more compact. These states have to be represented by a series of energy levels. The number of occupied energy levels is presumed to decrease as the structures become more defined during folding. Enzyme-catalyzed pathways involve the making and breaking of chemical bonds, and, hence, the usual transition-state theory can be applied directly. On the other hand, the pathway of protein folding consists of a series of conformational changes that conform to classical statistical mechanics. Since the proteins discussed here fold by clean first-order kinetics on a relatively slow time scale, we assume that these many conformational changes constitute a rapid series of preequilibria and that the slow first-order step represents the passing of the reaction over the highest energy species, which is the transition state for the reaction and occurs at a saddle point (28, 41). An equation analogous to that for transition state theory may be derived to describe this by using absolute rate theory—i.e., $k = \kappa\nu\exp(-\Delta G^{\ddagger}/RT)$, where $\Delta G^{\ddagger}$ is the mean difference in energy between the conformations at the saddle point of the reaction and the ground state, ν is a characteristic vibration frequency along the reaction coordinate at the saddle point, κ is a transmission coefficient, R is the gas constant, and T is the absolute temperature. Whereas the enzyme-kinetics analysis applies to the steady state, protein folding via intermediates occurs by a series of exponential steps. We can relate the time taken to fold in a multi-step reaction compared with a single step by using the following relationships. For a single-step reaction of rate constant k, the time taken for a fraction of $(e - 1)/e$ molecules to form products is $1/k$ (where $e = 2.718$), for a sequential two-step reaction of rate constants k_1 and k_2 is $1/k_1 + 1/k_2$, and for a multistep reaction of n steps, $\Sigma\, 1/k_i$ (see ref. 43).

In case A of Fig. 2, there is no intermediate and $G_{D_{phys}} \approx G_U$. There is a simple first-order folding reaction from D_{phys} to N with rate constant $k = \kappa\nu\exp(-\Delta G^{\ddagger}_A/RT)$. Case B is like case A, but $G_{D_{phys}} < G_U$ because stabilizing interactions are present in D_{phys}. The activation energy for the simple first-order kinetics is higher than that for case A by an amount $G_U - G_{D_{phys}}$ and so $k = \kappa\nu\exp[-(\Delta G^{\ddagger}_A + G_U - G_{D_{phys}})/RT]$. Case C is like case A, but with a high-energy intermediate I present ($G_{D_{phys}} < G_I$). This also obeys first-order kinetics with the same activation energy as in case A, such that the presence of the intermediate has no observed effect on the kinetics. Case D is like case C, but the intermediate is of low energy ($G_{D_{phys}} > G_I$). The reaction now has two distinct steps as D_{phys} progresses to N, the second one of which (k_2) has a higher activation energy than the single step for case C, where $k_2 = \kappa\nu\exp[-(\Delta G^{\ddagger}_A + G_{D_{phys}} - G_I)/RT]$. The time taken for a fraction of $(e - 1)/e$ molecules to fold from D_{phys} to N is: $\tau_A = \kappa^{-1}\nu^{-1}\exp(\Delta G^{\ddagger}_A/RT)$ for case A; $\tau_B = \kappa^{-1}\nu^{-1}\exp[(\Delta G G^{\ddagger}_A + G_U - G_{D_{phys}})/RT]$ for case B, such that $\tau_B > \tau_A$; $\tau_C = \kappa^{-1}\nu^{-1}\exp(\Delta G^{\ddagger}_A/RT)$ for case C; and $\tau_D > \kappa^{-1}\nu^{-1}\exp(\{\Delta G^{\ddagger}_A + G_{D_{phys}} - G_I\}/RT)$ for case D, such that $\tau_C = \tau_A < \tau_D$. Thus, protein folding is analogous to enzyme catalysis: the rate of folding drops when the collection of states that constitute D_{phys} is at a lower energy than those for U. Similarly, any intermediate between D_{phys} and the transition state that is of lower energy than D_{phys} also slows down the rate of folding since the time taken to fold is the sum of the times for each step. High energy intermediates do not slow down the rate of folding, provided their presence does not lead to states that are higher in energy than the transition state for the single-step reaction.

This has two important consequences. First, the accumulation of folding intermediates is unfavorable, as also argued by Creighton (16) and Sosnick *et al.* (17). Second, stable elements of structure in D_{phys} slow down folding since they lower, by definition, the energy levels of D_{phys}. It is intuitively obvious

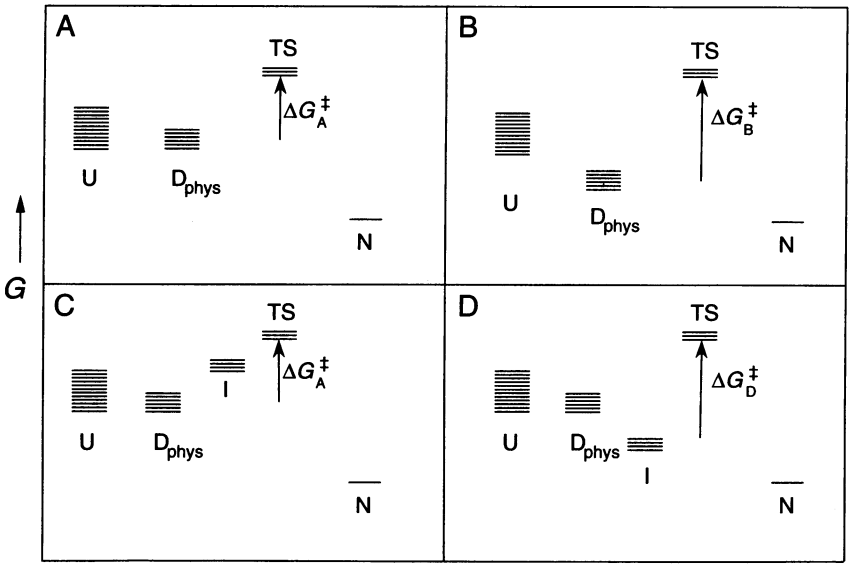

FIG. 2. Energy level diagrams for protein folding. The unfolded state (U) that is normally present under denaturing conditions is a family of conformations, some of which will have interactions not present in the final native structure and others that will. The denatured state under physiological conditions that favor folding (D_{phys}) also consists of many conformations. The conformations of the denatured states are drawn, for convenience, as stacked. Those denatured conformations that have some elements of native structure could be drawn more accurately as being on the reaction pathway and closer to the transition state. Case A; there is no intermediate and $G_{D_{phys}} \approx G_U$. Case B; like case A, but $G_{D_{phys}} < G_U$ because stabilizing interactions are present in D_{phys}. The activation energy for the simple first-order kinetics is higher than that for case A. Case C; like case A, but there is a high energy intermediate I present ($G_{D_{phys}} < G_I$). This also obeys first-order kinetics with the same activation energy as in case A, such that the presence of the intermediate has no observed effect on the kinetics. Case D; like case C, but the intermediate is low energy ($G_{D_{phys}} > G_I$). The reaction now has two distinct steps as D_{phys} progresses to N, the second one of which (k_2) has a higher activation energy than the single step for case C.

that nonnative interactions in D_{phys}, such as extensive hydrophobic interactions, will slow down folding. But, counter to intuition, the formation of stable elements of native structures in D_{phys} is also unfavorable in the sense that the protein has not evolved as far as possible for rapid folding. To illustrate why this is so, consider the example of adding helix-stabilizing residues to a protein that has an α-helix in its native structure and in its folding transition state, but not in its denatured state. This increases overall stability since the native structure of the protein will be stabilized relative to the denatured state. Further, if the protein folds by two-state kinetics (as in Fig. 2, case A), then stabilization of the helix in the transition state will also speed up the reaction as the energy barrier between the transition state and the denatured state decreases. Suppose that the helix becomes stabilized on the addition of several residues to such an extent that it becomes formed also in the denatured state. Any further addition of helix-stabilizing residues will not speed up the reaction since the energies of both the denatured and transition states will change in parallel. Similarly, the overall stability of the protein will not be increased. Thus, the extra stabilization energy is wasted. In quantitative terms, suppose that in D_{phys} the unfolded helix is in rapid equilibrium with its folded form and the microequilibrium constant for folding is k. The rate constant for folding (Fig. 1 *A* and *B*) is lowered by a factor $(1 + k)$, as is also the overall equilibrium constant for folding, from what it would be if the helix were completely unfolded. There is a window of stability for a local structural element: if too weak, it will destabilize the folded structure; if too stable, it will overstabilize the denatured state. This is one reason why it is favorable for a nucleation site to consist of a group of adjacent residues—e.g., part of an α-helix—which is not stable by itself but becomes stable by interacting with other residues elsewhere in the protein as it begins to fold. This means that the search for initiation sites in fragments of proteins or local elements of structure in intact proteins could frequently be futile.

The efficient nucleation–condensation mechanism of CI2, which conforms to simple two-state kinetics without the accumulation of an intermediate, fits precisely the criteria for rapid folding. Thus, CI2 appears to have evolved to fold rapidly. How general are these observations about the lack of stable structures in denatured states?

Lack of Stable Intermediates in Other Fast-Folding Small Proteins

Simple two-state kinetics of folding has now been found for the 56-residue IgG-binding domain of streptococcal protein G (20), the 62-residue Src homology 3 (SH3) domain of spectrin (21), the 86-residue acyl coenzyme A-binding protein (23), a modified form of cytochrome (17), an 80-residue fragment of the phage λ repressor (25), and a 67-residue fragment of CsbB (24). Lattice simulations show that fast folding can occur without a hydrophobic collapse (47).

Residual Structure in Denatured States and Protein Fragments

The detection of stable local elements of secondary structure—e.g., α-helices—in isolated fragments and the prediction of stable local elements in the denatured states of intact proteins have been used as evidence for sites of initiation of protein folding and for the order of events in pathways (48–52). But my analysis shows that the nucleation sites are best when they are only embryonic, and the structures do not accumulate in small fragments. This is indeed found in practice. Isolated fragments of α-helices or other elements of secondary structure in native proteins generally consist mainly of random coil with a small fraction of helix (53–58). Thirty percent α-helical conformation is considered impressive, but this would hardly alter the kinetics (only by a factor of 1.2). Isolated fragments of secondary structure are inherently weakly structured; therefore, little evolutionary pressure is required to keep them unstructured.

NMR analysis of the denatured state of the 110-residue barnase has shown only trace amounts of structure in the regions that correspond to the α-helices and β-sheet in the native structure (59). On the other hand, a fragment of the 434 repressor in 7 M urea contains a four-residue hydrophobic cluster, but this is in equilibrium with a considerable proportion of random coil (60).

The Folding of Larger Proteins

CI2 is a single module of protein structure and folds as a single cooperative unit. Proteins made of several modules have been thought to fold by the folding of individual modules via collapse mechanisms followed by the coalescing of the modules (26). According to the criteria for rapid folding, the optimal situation should be when coalescence coincides with the initial nucleation processes (a multinucleation mechanism). This must be difficult because so many residues must align simultaneously as complexity increases. For example, barnase, which is nearly twice the size of CI2 and has a modular structure (61), folds via a distinct folding intermediate at 25°C (13). On the other hand, at higher temperatures (M. T. Oliveberg, P. Dalby, and A.R.F., unpublished data), in the presence of denaturants (62), or at lower pH (41, 63), the folding intermediate becomes less stable, and the kinetics either tend to or actually become simple two state. Still larger proteins often fold very slowly *in vitro*. Lattice simulations of larger proteins also show transitions to folding in separate domains by multinucleation pathways (36).

The folding of larger proteins *in vivo* is believed to invoke molecular chaperones, one of whose roles is thought to be the unfolding of misfolded intermediates. Interestingly, small proteins fold well *in vitro* without the mediation of molecular chaperones and probably do not use chaperone-mediated folding *in vivo*. The mechanism of folding of CI2 suggests that small proteins have evolved to fold fast.

Conclusions

(*i*) A nucleation mechanism or variations thereon in which the nucleation site occurs only flickeringly in the denatured state and no folding intermediate accumulates is an efficient folding pathway of a small protein.

(*ii*) Evolutionary pressure therefore opposes the accumulation of nucleation sites in the denatured state. Sites should become stable only after interacting with other parts of the protein. This discourages searches for structure in isolated fragments by experimentalists or for initiation sites in intact proteins by theoreticians using methods that disregard long-distance interactions.

(*iii*) Evolutionary pressure should minimize nonnative hydrophobic interactions in denatured states since such interactions lower the energies of the denatured states.

Note Added in Proof. A lattice simulation has now also shown that optimal conditions for folding are achieved when the contributions to stability from local interactions are small and that further increases in the local interactions in proteins that are optimally evolved for folding decreases their foldability (66).

1. Levinthal, C. (1968) *J. Chim. Phys.* **85,** 44–45.
2. Ptitsyn, O. B. (1973) *Dokl. Acak. Nauk. SSR* **210,** 1213–1215.
3. Ptitsyn, O. B. (1994) *Protein Eng.* **7,** 593–596.
4. Karplus, M. & Weaver, D. C. (1976) *Nature (London)* **260,** 404–406.
5. Karplus, M. & Weaver, D. L. (1994) *Protein Sci.* **3,** 650–668.

6. Kim, P. S. & Baldwin, R. L. (1982) *Annu. Rev. Biochem.* **51,** 459–489.
7. Kim, P. S. & Baldwin, R. L. (1990) *Annu. Rev. Biochem.* **59,** 631–660.
8. Dill, K. (1985) *Biochemistry* **24,** 1501–1509.
9. Dill, K. A., Bromberg, S., Yue, K. Z., Fiebig, K. M., Yee, D. P., Thomas, P. D. & Chan, H. S. (1995) *Protein Sci.* **4,** 561–602.
10. Kuwajima, K. (1989) *Proteins Struct. Funct. Genet.* **6,** 87–103.
11. Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Zerovnik, E. & Razgulyaev, O. I. (1990) *FEBS Lett.* **262,** 20–24.
12. Bychkova, V. E. & Ptitsyn, O. B. (1993) *Biofizika* **38,** 58–66.
13. Fersht, A. R. (1993) *FEBS Lett.* **325,** 5–16.
14. Fersht, A. R. (1974) *Proc. R. Soc. London B* **187,** 397–407.
15. Go, N. (1983) *Annu. Rev. Biophys. Bioeng.* **12,** 183–210.
16. Creighton, T. E. (1994) *Nat. Struct. Biol.* **1,** 135–138.
17. Sosnick, T. R., Mayne, L., Hiller, R. & Englander, S. W. (1994) *Nat. Struct. Biol.* **1,** 149–156.
18. Jackson, S. E. & Fersht, A. R. (1991) *Biochemistry* **30,** 10428–10435.
19. Jackson, S. E. & Fersht, A. R. (1991) *Biochemistry* **30,** 10436–10443.
20. Alexander, P., Orban, J. & Bryan, P. (1992) *Biochemistry* **31,** 7243–7248.
21. Viguera, A. R., Martinez, J. C., Filimonov, V. V., Mateo, P. L. & Serrano, L. (1994) *Biochemistry* **33,** 2142–2150.
22. Kuszewski, J., Clore, G. M. & Gronenborn, A. M. (1994) *Protein Sci.* **3,** 1945–1952.
23. Kragelund, B. B., Robinson, C. V., Knudsen, J., Dobson, C. M. & Poulsen, F. M. (1995) *Biochemistry* **34,** 7217–7224.
24. Schindler, T. S., Herrier, M., Marahiel, M. A. & Schmid, F. X. (1995) *Nat. Struct. Biol.* **2,** 663–673.
25. Huang, G. S. & Oas, T. G. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 6878–6882.
26. Otzen, D. E., Itzhaki, L. S., Elmasry, N. F., Jackson, S. E. & Fersht, A. R. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 10422–10425.
27. Matouschek, A., Kellis, J. T., Jr., Serrano, L. & Fersht, A. R. (1989) *Nature (London)* **342,** 122–126.
28. Fersht, A. R., Matouschek, A. & Serrano, L. (1992) *J. Mol. Biol.* **224,** 771–782.
29. Fersht, A. R. (1995) *Philos. Trans. R. Soc. London B* **348,** 11–15.
30. Shakhnovich, E. I., Farztdinov, G., Gutin, A. M. & Karplus, M. (1991) *Phys. Rev. Lett.* **67,** 1665–1668.
31. Skolnick, J. & Kolinski, A. (1991) *J. Mol. Biol.* **221,** 499–531.
32. Miller, R., Danko, C. A., Fasolka, M. J., Balazs, A. C., Chan, H. & Dill, K. A. (1992) *J. Chem. Phys.* **96,** 768–780.
33. Camacho, C. J. & Thirumalai, D. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 6369–6372.
34. Abkevich, V. I., Gutin, A. M. & Shakhnovich, E. I. (1994) *Biochemistry* **33,** 10026–10036.
35. Sali, A., Shakhnovich, E. & Karplus, M. (1994) *J. Mol. Biol.* **235,** 1614–1636.
36. Abkevich, V. I., Gutin, A. M. & Shakhnovich, E. I. (1995) *Protein Sci.* **4,** 1167–1177.
37. Onuchic, J. N., Wolynes, P. G., Lutheyschulten, Z. & Socci, N. D. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 3626–3630.
38. Yue, K., Fiebig, K. M., Thomas, P. D., Chan, H. S., Shakhnovich, E. I. & Dill, K. A. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 325–329.
39. Guo, Z. Y. & Thirumalai, D. (1995) *Biopolymers* **36,** 83–102.
40. Gay, G. D., Ruizsanz, J., Neira, J. L., Itzhaki, L. S. & Fersht, A. R. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 3683–3686.
41. Oliveberg, M., Tan, Y.-J. & Fersht, A. R. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 8926–8929.
42. Wetlaufer, D. B. (1973) *Proc. Natl. Acad. Sci. USA* **70,** 697–701.
43. Fersht, A. R. (1985) *Enzyme Structure and Mechanism* (Freeman, New York), 2nd Ed.
44. Haldane, J. B. S. (1930) *Enzymes* (Longmans Green, London).
45. Pauling, L. (1946) *Chem. Eng. News* **24,** 1375–1377.
46. Serrano, L., Day, A. G. & Fersht, A. R. (1993) *J. Mol. Biol.* **233,** 305–312.
47. Gutin, A. M., Abkevich, V. I. & Shakhnovich, E. I. (1995) *Biochemistry* **34,** 3066–3076.
48. Wright, P. E., Dyson, H. J. & Lerner, R. A. (1988) *Biochemistry* **27,** 7167–7175.
49. Dyson, H. J. & Wright, P. E. (1991) *Annu. Rev. Biophys. Biophys. Chem.* **20,** 519–538.
50. Chelvanayagam, G., Reich, Z., Bringas, R. & Argos, P. (1992) *J. Mol. Biol.* **227,** 901–916.
51. Rooman, M. J., Kocher, J. P. A. & Wodak, S. J. (1992) *Biochemistry* **31,** 10226–10238.
52. Rooman, M. J. & Wodak, S. J. (1992) *Biochemistry* **31,** 10239–10249.
53. Waltho, J. P., Feher, V. A., Merutka, G., Dyson, H. J. & Wright, P. E. (1993) *Biochemistry* **32,** 6337–6347.
54. Shin, H. C., Merutka, G., Waltho, J. P., Wright, P. E. & Dyson, H. J. (1993) *Biochemistry* **32,** 6348–6355.
55. Shin, H. C., Merutka, G., Waltho, J. P., Tennant, L. L., Dyson, H. J. & Wright, P. E. (1993) *Biochemistry* **32,** 6356–6364.
56. Blanco, F. J. & Serrano, L. (1995) *Eur. J. Biochem.* **230,** 634–649.
57. Munoz, V., Serrano, L., Jimenez, M. A. & Rico, M. (1995) *J. Mol. Biol.* **247,** 648–669.
58. Sancho, J., Neira, J. L. & Fersht, A. R. (1992) *J. Mol. Biol.* **224,** 749–758.
59. Arcus, V. L., Vuilleumier, S., Freund, S. M. V., Bycroft, M. & Fersht, A. R. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 9412–9416.
60. Neri, D., Billeter, M., Wider, G. & Wuthrich, K. (1992) *Science* **257,** 1559–1563.
61. Yanagawa, H., Yoshida, K., Torigoe, C., Park, J. S., Sato, K., Shirai, T. & Go, M. (1993) *J. Biol. Chem.* **268,** 5861–5865.
62. Matouschek, A., Kellis, J. T., Jr., Serrano, L., Bycroft, M. & Fersht, A. R. (1990) *Nature (London)* **346,** 440–445.
63. Oliveberg, M. T. & Fersht, A. R. (1995) *Biochemistry*, in press.
64. Itzhaki, L. S., Neira, J. L., Ruiz-Sanz, J., Gay, G. D. & Fersht, A. R. (1995) *J. Mol. Biol.*, in press.
65. Tan, Y. J., Oliveberg, M., Davis, B. & Fersht, A. R. (1995) *J. Mol. Biol.*, in press.
66. Govindarajan, S. & Goldstein, R. A. (1995) *Proteins Struct. Funct. Genet.* **22,** 413–418.